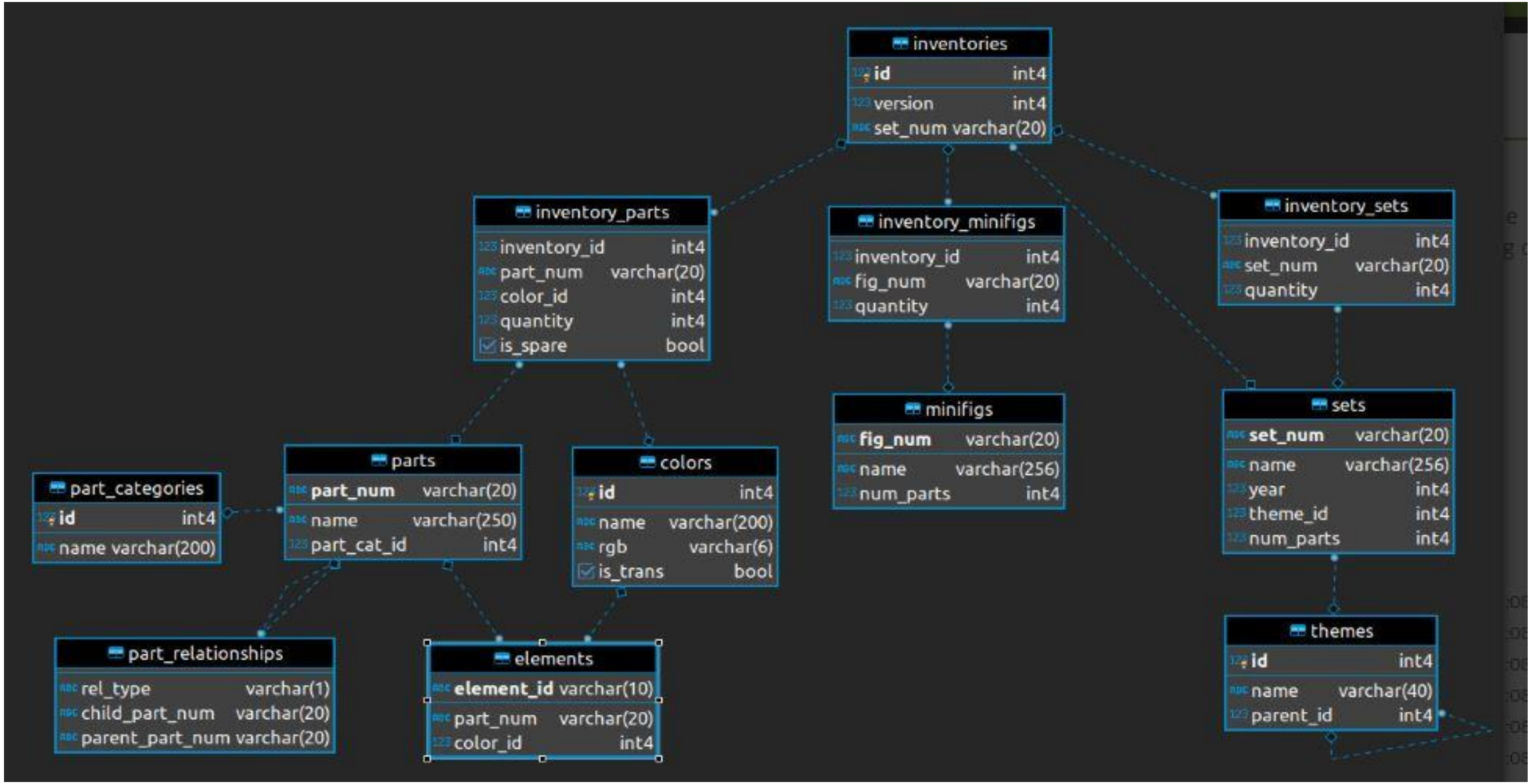


LEGO Catalog Dataset- Data Analysis Using SQL (MSSQL) Source-https://rebrickable.com/downloads/

```sql
--What is the number of parts/per
create view dbo.analytics_main as (

select s.set_num, s.name as set_nam,s.year,s.theme_id, cast(s.num_parts as numeric) num_parts,
t.name as theme_name,
t.parent_id,p.name as parent_theme_name
from dbo.sets s
left join [dbo].[themes] t
   on s.theme_id=t.id
   left join [dbo].[themes] p
   on t.parent_id=p.id)
(select *
   from dbo.analytics_main)

WITH cte_ThemeCount
AS (
select cast(num_parts as numeric) num_parts,[dbo].[themes].name,parent_id
from [dbo].[sets]
left join [dbo].[themes]
on [dbo].[sets].[theme_id]=[dbo].[themes].id)
SELECT sum(num_parts) as Total_Number_of_parts,name FROM cte_ThemeCount
where parent_id is not null
group by name
```

What is the number of parts per theme

Results | Messages

	Total_Number_of_parts	name
1	107879	Ultimate Collector Series
2	81043	Creator 3-in-1
3	79954	Creator Expert
4	51149	Classic Town
5	48966	Avengers
6	39998	Basic
7	35315	Police
8	31030	Serious Play
9	29415	Batman
10	28432	Bricks & More
11	26680	Town
12	25057	Christmas
13	24448	Technic
14	23347	Spider-Man
15	22459	Duplo and Explore
16	21483	The LEGO Ninjago M...
17	20970	Disney Princess

Query executed successfully.

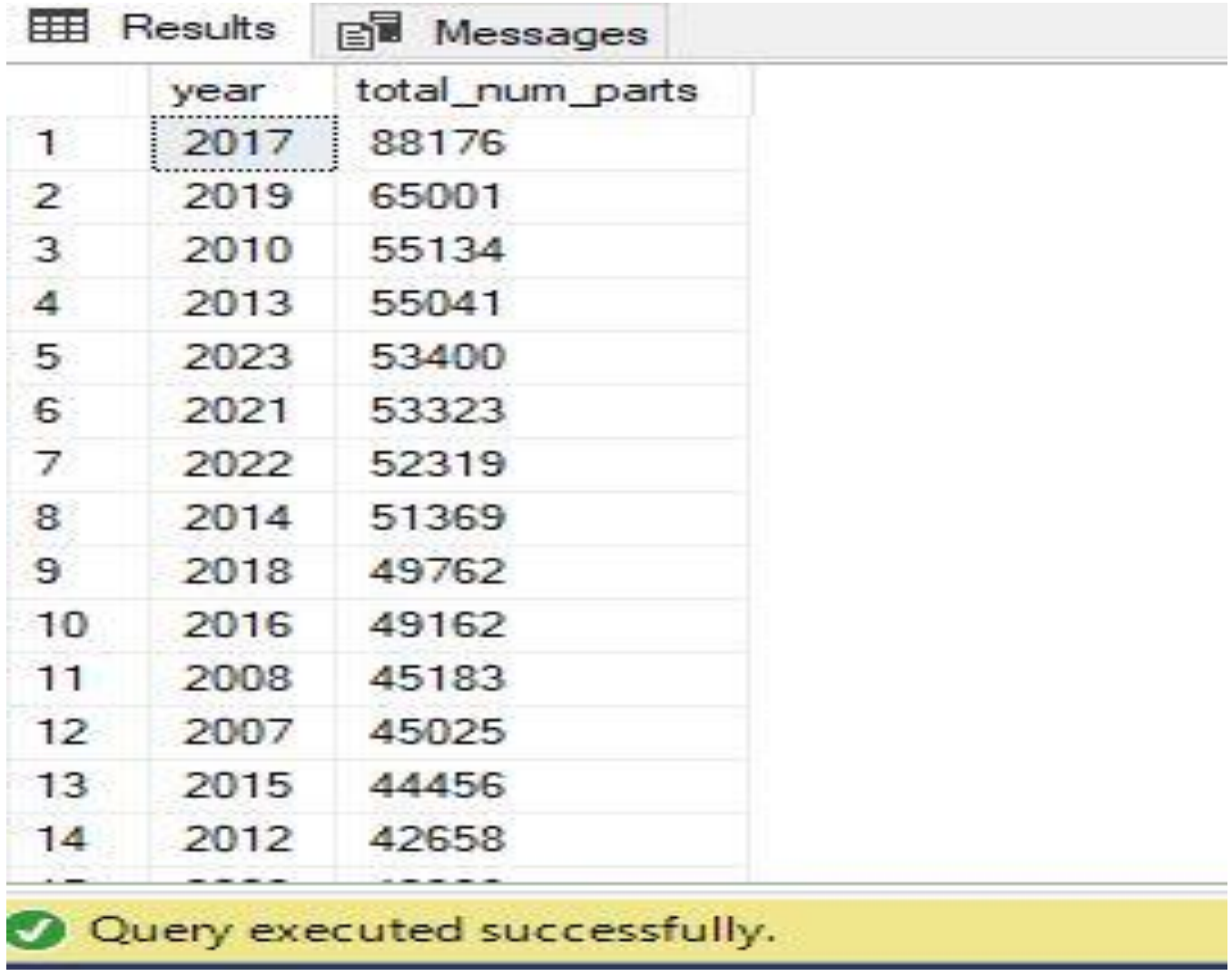
Query showed total number of parts/theme. “Ultimate Collector Series Theme adds up to maximum”

```sql
---What is the number of parts per year
select year,sum(num_parts) as total_num_parts
from [dbo].[analytics_main]
where parent_theme_name is not null
group by year
order by 2 desc
```

```sql
create view dbo.analytics_main as (

select s.set_num, s.name as set_nam,s.year,s.theme_id, cast(s.num_parts as numeric) num_parts,
t.name as theme_name,
t.parent_id,p.name as parent_theme_name
from dbo.sets s
left join [dbo].[themes] t
   on s.theme_id=t.id
   left join [dbo].[themes] p
   on t.parent_id=p.id)
```

What is the number of parts per year


Results Messages

	year	total_num_parts
1	2017	88176
2	2019	65001
3	2010	55134
4	2013	55041
5	2023	53400
6	2021	53323
7	2022	52319
8	2014	51369
9	2018	49762
10	2016	49162
11	2008	45183
12	2007	45025
13	2015	44456
14	2012	42658

Query executed successfully.

Query resulted total number of parts grouped by each Year.

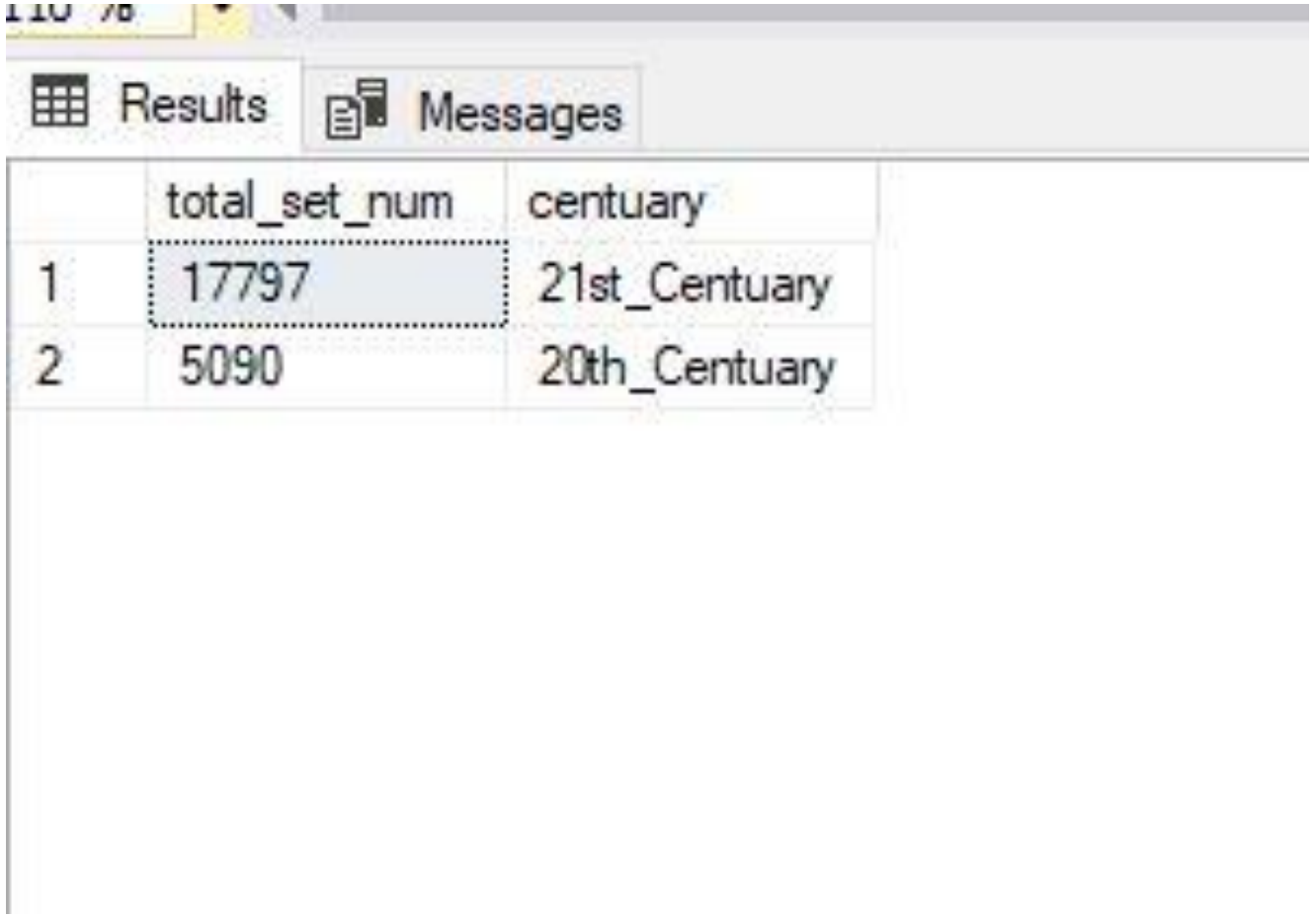
```sql
-----How may sets where created in each centuary in the dataset

select count(set_num) as total_set_num,centuary
from dbo.analytics_main
--where parent_theme_name is not null
group by centuary
order by 1 desc
```

```sql
create view dbo.analytics_main as (

select s.set_num, s.name as set_nam,s.year,s.theme_id, cast(s.num_parts as numeric) num_parts,
t.name as theme_name,
t.parent_id,p.name as parent_theme_name
from dbo.sets s
left join [dbo].[themes] t
   on s.theme_id=t.id
   left join [dbo].[themes] p
   on t.parent_id=p.id)
```

Query to find how many sets where created in each century in given dataset



Query resulted total number of theme’ sets by century.

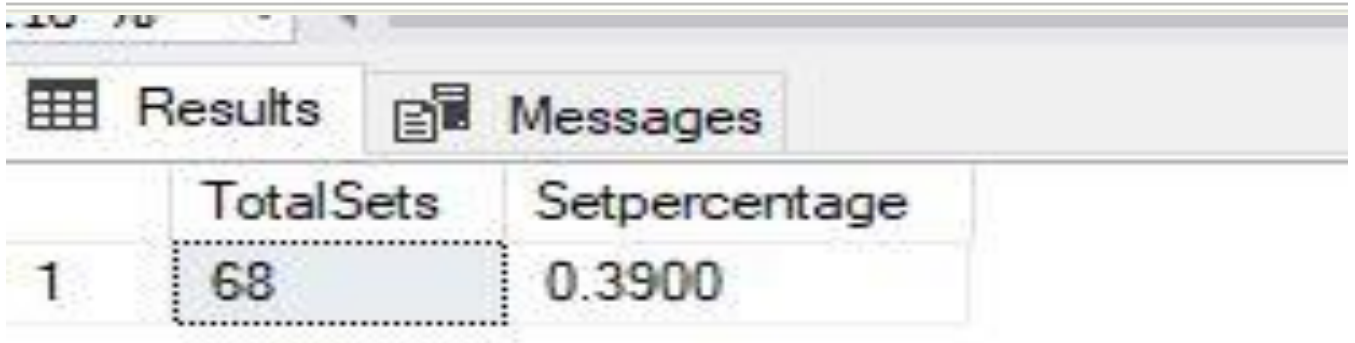
```sql
---what percentage of sets ever released in the 21st centuary were Trains Themed
WITH CTE_Percentage AS(
Select count(set_num)as total_set_num,centuary,theme_name
from [dbo].[analytics_main]
where centuary = '21st_Centuary'
group by theme_name,Centuary)

select sum(total_set_num) as TotalSets,sum(sets_percentage) as Setpercentage
from(
select Centuary, theme_name,total_set_num,
sum(total_set_num) OVER()as total, Cast((1.00*total_set_num / sum(total_set_num) OVER()) as decimal(5,4))*100 sets_percentage
from CTE_Percentage)h
where theme_name like '%train%'
```

Query to find percentage of sets ever released in 21st century and have theme named “Train”

Results | Messages

	Centuary	theme_name	total_set_num	total	sets_percentage
1	21st_Centuary	Train	3	68	4.4100
2	21st_Centuary	Trains	65	68	95.5900



Out of total released themes, only 0.39%(<1%) of were named Train

```sql
---What was the popular theme by year in terms of sets released in the 21st centuary
WITH CTE_Percentage AS(
Select count(set_num)as total_set_num,centuary,theme_name,year
from [dbo].[analytics_main]
where centuary = '21st_Centuary'
group by theme_name,Centuary,year)

Select *
from(
Select count(set_num) as TotalSetNum,centuary,theme_name,
    year,ROW_NUMBER() OVER(partition by year order by count(set_num) desc) rownum
from [dbo].[analytics_main]
where centuary = '21st_Centuary'
group by theme_name,Centuary,year) g
where rownum =1
order by TotalSetNum desc
```

What was the popular theme by year in terms of sets released in 21st Century

Results | Messages

	TotalSetNum	centuary	theme_name	year	rownum
1	156	21st_Centuary	Gear	2021	1
2	118	21st_Centuary	Technic	2014	1
3	69	21st_Centuary	Star Wars	2015	1
4	69	21st_Centuary	Star Wars	2018	1
5	64	21st_Centuary	The LEGO Movie II	2019	1
6	64	21st_Centuary	Star Wars	2017	1
7	64	21st_Centuary	Technic	2002	1
8	63	21st_Centuary	Bionicle	2003	1
9	63	21st_Centuary	Star Wars	2016	1
10	60	21st_Centuary	Gear	2022	1
11	59	21st_Centuary	Gear	2023	1
12	59	21st_Centuary	Friends	2020	1
13	55	21st_Centuary	Bionicle	2006	1
14	51	21st_Centuary	Bionicle	2004	1
15	47	21st_Centuary	Ninjago	2012	1
16	45	21st_Centuary	Bionicle	2001	1
17	45	21st_Centuary	Friends	2024	1

Query executed successfully.

Out of total released themes, “Gear” Theme is the most popular in 21st century

```sql
Create view itsdef as (
select c.id,c.name as color_name,inv.inventory_id,inv.quantity,parts.part_num,parts.name
from colors c
left join inventory_parts inv
on c.id=inv.color_id
left join parts
on inv.part_num=parts.part_num)

select max(QtySum) as result,color_name from(
select sum(quantity) as QtySum,id,color_name
from itsdef
group by id,color_name)main
group by color_name
order by result desc
```

what is the most produced colour of LEGO ever in terms of quantity of parts?

Results | Messages

	result	color_name
1	789844	Black
2	487181	Light Bluish Gray
3	473741	White
4	346867	Dark Bluish Gray
5	306139	Red
6	204812	Yellow
7	204498	Blue
8	168569	Tan
9	145464	Reddish Brown
10	105925	Light Gray
11	79260	Green
12	55669	Dark Tan
13	47851	Orange
14	46719	Dark Blue

Query executed successfully.

It is clear the Black colour has highest quantity per parts